Catalyst 9000シリーズスイッチでのSSDPベストプラクティスの実装

内容

概要

このドキュメントでは、Catalyst 9000シリーズスイッチでSimple Service Discovery Protocol(SSDP)パケットをドロップまたは制限するためのベストプラクティス設定について説明します。

前提条件

要件

次の項目に関する知識があることが推奨されます。

- Protocol Independent Multicast(PIM)の動作
- お客様の環境に固有のSSDPの使用方法

使用するコンポーネント

このドキュメントの情報は、次のソフトウェアとハードウェアのバージョンに基づいています。

- Cisco Catalyst 9200
- Cisco Catalyst 9300
- Cisco Catalyst 9400
- Cisco Catalyst 9500
- Cisco Catalyst 9600

このドキュメントの情報は、特定のラボ環境にあるデバイスに基づいて作成されました。このドキュメントで使用するすべてのデバイスは、初期（デフォルト）設定の状態から起動しています。本稼働中のネットワークでは、各コマンドによって起こる可能性がある影響を十分確認してください。

エンタープライズ環境におけるSSDPリスクの理解

一般に、ラップトップや携帯電話などのエンドユーザデバイスは、SSDPプロトコルを使用するユニバーサルプラグアンドプレイ(UPnP)機能を自動的にアドバタイズします。クライアントは、IPアドレス239.255.255.250にマルチキャストアドバタイズメントパケットを送信します。これらのアドバタイズメントは、多くの場合、存続可能時間(TTL)が1で送信され、マルチキャストパケットを生成したホストのローカルサブネットを超えることはありません。ネットワーク上の他のデバイスのアドバタイズメントを受信するために、エンドポイントはIGMPメンバシップレポートを239.255.255.250アドレスに送信します。これは、他のマルチキャスト送信元からこのIPアドレスに送信されたマルチキャストトラフィックもこのクライアントに転送する必要があることをネットワークに通知します。

すべてのエンドポイントがソースとして機能し、このグループの対象レシーバとして機能する数百または数千のエンドポイントを含むエンタープライズ環境では、このクライアントアクティビティをオフにしておくとネットワークデバイスに過大な負荷をかけやすく、ネットワークリソースが使い果たされるとサービスが停止する可能性があります。

この枯渇は、主に次の2つの方法のいずれかで発生します。

1. 二次プロトコル障害を引き起こすハードウェアリソースの枯渇
2. 分散型サービス拒否(DDoS)攻撃として使用されるSSDPからのインターフェイスとプラットフォームの帯域幅枯渇。

このドキュメントでは詳しく説明していませんが、SSDPのオープンな性質により、攻撃者がこのサービスを有効にしたクライアントのグループに巧妙に細工されたパケットを送信して、1つまたは複数の宛先ホストに大きな応答を送信する可能性があることに注意してください。また、大量の発信インターフェイスステートが作成されると、スイッチは特定用途向け集積回路(ASIC)内の発信インターフェイスごとに各フレームのコピーを1つ作成する必要があるため、少量のマルチキャストトラフィックによってスイッチのパフォーマンス容量に大きな負荷がかかる可能性があります。 発信インターフェイスは、20以上のインターフェイスがキャパシティの問題やパケット損失のリスクが高いことを示します。

ハードウェアリソース枯渇の症状

Catalyst 9000シリーズスイッチでは、リソースが使い果たされると、「fman_fp_image」または「FMFP」と表示されるsyslogが出力されます。これらのエラーの一部またはすべてが、スイッチでリソースの枯渇が発生し、さらに調査が必要になった場合に出力される可能性があります。

これらは、リソースの枯渇時に発生する一般的なエラーの一部ですが、包括的なリストではありません。

図 1：スイッチのリソース枯渇の証拠となる、最も一般的なエラーのサンプル

```
%FMFP-3-OBJ_DWNLD_TO_DP_STUCK: R0/0: fman_fp_image: AOM download to Data Plane is stuck for more
than 1800 seconds for <object details>
%FMFP-3-OBJ_DWNLD_TO_DP_RESUME: R0/0: fman_fp_image: AOM download of objects to Data Plane is
back to normal
%FMFP_QOS-6-QOS_STATS_STALLED: R0/0: fman_fp_image: statistics stalled
%FMFP-3-OBJ_DWNLD_TO_DP_FAILED: R0/0: fman_fp_image: adj <hex>, Flags None download to DP failed
%FMFP-3-OBJ_DWNLD_TO_DP_FAILED: R0/0: fman_fp_image: adj <hex>, Flags Midchain download to DP
failed
%FED_L3M_ERRMSG-3-RSRC_ERR: Switch <num> R0/0: fed: Failed to allocate hardware resource for
group <address> - rc:<number or error>
%FED_L3_ERRMSG-3-RSRC_ERR: Chassis <num> R0/0: fed: Failed to allocate hardware resource for adj
```

entry due to hardware resource exhaustion - rc:<number or error>

SSDPによるハードウェアリソース枯渇の確認

すべてのCatalyst 9000シリーズスイッチは、特別なASICを使用して、パケットルーティングの大部分を高スループットで実行します。これらのASICは、容量が限られているさまざまなテーブルと内部リソースを活用します。SSDPクライアントは、共通のマルチキャストグループの送信元と受信者の両方として機能するため、ハードウェアは、これらの制限されたリソースを使用して、他の理由(TTL 1)でパケットが送受信されない場合でも、パケットが通過するパスをハードウェアでプログラムする必要があります。 ハードウェアリソースが使い果たされると、SSDPとの関係に関係なく、どのグループについても新しいアップデートや追加をインストールできなくなります。インストールされていない多数のSSDP更新（状態の変更）もソフトウェアでキューに入れられます。これにより、非マルチキャストトラフィックのハードウェア更新が中断または失敗する可能性があり、ユーザトラフィックに影響を与え、ネットワークの停止を引き起こします。

このドキュメントは、ネットワークがPIMで設定され、既知のSSDPグループアドレスに対してレイヤ3マルチキャスト状態である場合にのみ関連します。この基準を確認するには、次のコマンドを実行します "**show ip mroute 239.255.255.250**"（必要に応じてvrfステートメントを追加します）。 グループ239.255.255.250はSSDPプロトコルに固有です。

コマンド出力に多数の発信インターフェイスが含まれているか、この特定のグループに対して多数の固有のソースが含まれている場合、システムおよびネットワークがSSDPによる停止に対して脆弱であることを示しています。発信インターフェイスと一意の送信元の数が多いほど、サービスに影響する可能性が高くなります。

図 2： 出力例 "show ip mroute 239.255.255.250" コマンドを発行します。

```
Switch#show ip mroute 239.255.255.250
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,
       L - Local, P - Pruned, R - RP-bit set, F - Register flag,
       T - SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet,
       X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement,
       U - URD, I - Received Source Specific Host Report,
       Z - Multicast Tunnel, z - MDT-data group sender,
       Y - Joined MDT-data group, y - Sending to MDT-data group,
       G - Received BGP C-Mroute, g - Sent BGP C-Mroute,
       N - Received BGP Shared-Tree Prune, n - BGP C-Mroute suppressed,
       Q - Received BGP S-A Route, q - Sent BGP S-A Route,
       V - RD & Vector, v - Vector, p - PIM Joins on route,
       x - VxLAN group
Outgoing interface flags: H - Hardware switched, A - Assert winner, p - PIM Join
Timers: Uptime/Expires
Interface state: Interface, Next-Hop or VCD, State/Mode

(*, 239.255.255.250), 00:08:35/stopped, RP 10.0.0.1, flags: SJC
  Incoming interface: GigabitEthernet0/0/1.40, RPF nbr 10.0.0.1
  Outgoing interface list:
    GigabitEthernet0/0/1.101, Forward/Sparse, 00:08:35/00:02:40
    GigabitEthernet0/0/1.102, Forward/Sparse, 00:08:35/00:02:38
    GigabitEthernet0/0/1.100, Forward/Sparse, 00:08:35/00:02:39

(10.1.1.2, 239.255.255.250), 00:01:40/00:01:19, flags: T
  Incoming interface: Null, RPF nbr 0.0.0.0
  Outgoing interface list:
    GigabitEthernet0/0/1.40, Forward/Sparse, 00:01:40/00:01:40, A
```

```
    GigabitEthernet0/0/1.100, Forward/Sparse, 00:01:40/00:02:39
    GigabitEthernet0/0/1.102, Forward/Sparse, 00:01:40/00:02:38
    GigabitEthernet0/0/1.101, Forward/Sparse, 00:01:40/00:02:40

(10.1.1.3, 239.255.255.250), 00:02:03/00:00:56, flags: JT
  Incoming interface: GigabitEthernet0/0/1.40, RPF nbr 10.1.1.1
  Outgoing interface list:
    GigabitEthernet0/0/1.100, Forward/Sparse, 00:02:03/00:02:39
    GigabitEthernet0/0/1.102, Forward/Sparse, 00:02:03/00:02:38
    GigabitEthernet0/0/1.101, Forward/Sparse, 00:02:03/00:02:40

(10.1.1.4, 239.255.255.250), 00:08:35/00:02:32, flags: T
  Incoming interface: GigabitEthernet0/0/1.40, RPF nbr 10.1.1.1
  Outgoing interface list:
    GigabitEthernet0/0/1.100, Forward/Sparse, 00:08:35/00:02:39
    GigabitEthernet0/0/1.102, Forward/Sparse, 00:08:35/00:02:38
    GigabitEthernet0/0/1.101, Forward/Sparse, 00:08:35/00:02:40, A
```

SSDPが特定の目的で使用されない限り、この出力は空であるか、発信インターフェイスの数が少ないか、一意の送信元の数が少ないかのいずれかになります。これは、リソースの枯渇やサービスへの影響を防ぐためです。

多数のマルチキャストグループが見られる場合は、コマンド「**show platform software object-manager fp active statistics**」または「**show platform software object-manager fp switch active statistics**」を使用して、ハードウェアリソースが使い果たされているかどうかを確認できます。

> **注**：このコマンドは、マルチキャストトラフィックによってトリガーされるリソース枯渇に固有のものではなく、他の問題によってこれらの値がゼロ以外になる可能性があります。

図 3： 出力 **"show platform software object-manager fp active statistics"**問題状態にある

```
Switch#show platform software object-manager fp active statistics
Forwarding Manager Asynchronous Object Manager Statistics
Object update: Pending-issue: 109058, Pending-acknowledgement: 76928   <-- Pending-issue is very high, this
Batch begin:   Pending-issue: 0, Pending-acknowledgement: 0                  is not expected.
Batch end:     Pending-issue: 0, Pending-acknowledgement: 0
Command:       Pending-acknowledgement: 0
Total-objects: 304085
Stale-objects: 0
Resolve-objects: 0
Childless-delete-objects: 530
Error-objects: 1098

Paused-types: 127
```

図3の出力は、リソースが枯渇したスイッチの症状を示しています。通常の動作中には予想されないコマンド出力行がいくつかあります。

- Pending-issue:この値はゼロまたは近い値になります。これがコマンドの複数の繰り返しにわたって大きなゼロ以外の値のままである場合、これはリソースの枯渇の兆候です
- Pending-acknowledgement:この値はゼロまたは近い値になります。これがコマンドの複数の繰り返しにわたって大きなゼロ以外の値のままである場合、これはリソースの枯渇の兆候です
- Childless-delete-objects:この値は、ゼロまたは近い値になります。10以上の値は予期されて

いません。
- エラーオブジェクト：この値は、ゼロまたは近い値になります。10以上の値は予期されていません。

「pending-issue」カウンタまたは「pending-acknowledgement」カウンタが大量に存在する状態では、ハードウェアが誤ってプログラムされるリスクが高くなります。誤ってプログラムされたハードウェアは、ユニキャストおよびマルチキャストトラフィックの停止の一般的な原因です。

"**show platform hardware fed switch active fwd-asic resource utilization**" or in some models "**show platform hardware fed active fwd-asic resource utilization**" を使用して、ASICで使用されている有限のリソースの一部を調べ、内部リソースが使い果たされているかどうかを確認できます。

図 4： 出力例"show platform hardware fed active fwd-asic resource utilization"ほとんど使い果たされそうな一資源を使って

```
Switch#show platform hardware fed active fwd-asic resource utilization
Resource Info for ASIC Instance: 0
Resource Name           Allocated Free
----------------------------------------
RSC_DI                     3822       38076
RSC_FAST_DI                   0         192
RSC_RIET_0                    1        1024
RSC_RIET_1                    0         512
RSC_RIET_2                    0         512
RSC_RIET_3                    0         512
RSC_RIET_4                    0         512
RSC_RIET_5                    0         512
RSC_RIET_6                    0         256
RSC_RIET_7                    0         255
RSC_VLAN_LE                 116        3976
RSC_L3IF_LE                 116        3907
RIM_RSC_DGT                   1         255
RSC_VPN_PREFIX_ID             1       32768
RSC_LABEL_STACK_ID            1       65536
RSC_RI                     7358       82730
RSC_LI_RI                     0         129
RSC_PORT_LE_RI                0        2048
RSC_PORT_LE                   0        1827
RSC_RI_REP                10635      120437
RSC_SI                    11842      119072
RSC_SI_IND                    1         255
RSC_SI_STATS               3550       45602
RSC_RCP1_FID                  1        1023
RSC_RCP2_FID                  1        1023
RSC_RCP3_FID                  1        1023
RSC_RCP4_FID                  1        1023
RSC_LV1_ECR                   1          63
RSC_LV2_ECR                   3         253
RSC_ENH_ECR                   1           0
RSC_RPF_MATCH                12        1012
RSC_PLC                       1        2047
RSC_PLC_PF                    1         255
RSC_MTU_INDEX                 6         250
RSC_EGR_REDIRECT_INDEX        2        2046
RSC_RIL_INDEX 131065 7 <-- Free entries extremely low, this is not expected.
RSC_SIF                       1        1023
RSC_GROUP_LE                  1        1023
RSC_RI_REP_LOCAL              1           0
RSC_EXT_SI                  512       65024
```

図4では、「RSC_RIL_INDEX」の値は、使用中のエントリが131065個あり、空きがあるのは7個のみであることを示しています。このリソースは、多数の一意なSSDPグループによって消費されます。SSDPに固有ではありませんが、空きエントリ数が少なく、割り当て済みエントリ数が多いリソースは、スイッチがキャパシティの問題に近づいている兆候であり、調査する必要があります。

"**show platform hardware fed switch active fwd-asic resource tcam utilization**" or on some models "**show platform hardware fed active fwd-asic resource tcam utilization**" ASICごとのリソースごとの使用率の内訳を表示するために使用できます。SSDP枯渇のもう1つのシグニチャは、「L3マルチキャストエントリ」の「Used Values」列が「Max Values」に近い、または「Max Values」に達することです。

図 5： 出力例"show platform hardware fed active fwd-asic resource tcam utilization"正常な動作で

```
Switch#show platform hardware fed active fwd-asic resource tcam utilization
CAM Utilization for ASIC  [0]
 Table                                              Max Values        Used Values
 ------------------------------------------------------------------------------
 Unicast MAC addresses                               32768/768       6160/21
 L3 Multicast entries                                32768/768       3544/8      <-- Normal
Utilization, not near Max Values
 L2 Multicast entries                                 2304            181        <-- Normal
Utilization, not near Max Values
 Directly or indirectly connected routes            212992/1536     11903/39
 Input Ipv4 QoS Access Control Entries                5632            17
 Input Non Ipv4 QoS Access Control Entries            2560            36
 Output Ipv4 QoS Access Control Entries               6144            13
 Output Non Ipv4 QoS Access Control Entries           2048            27
 Input Ipv4 Security Access Control Entries           7168            12
 Input Non Ipv4 Security Access Control Entries       5120            76
 Output Ipv4 Security Access Control Entries          7168            11
 Output Non Ipv4 Security Access Control Entries      8192            27
 Ingress Netflow ACEs                                 1024             8
 Policy Based Routing ACEs                            3072            20
 Egress Netflow ACEs                                  1024             8
 Flow SPAN ACEs                                        512             5
 Flow Egress SPAN ACEs                                 512             8
 Control Plane Entries                                1024           235
 Tunnels                                              2816            26
 Lisp Instance Mapping Entries                         512             3
 Input Security Associations                           512             4
 SGT_DGT                                             32768/768         0/1
 CLIENT_LE                                            8192/512         0/0
 INPUT_GROUP_LE                                       1024             0
 OUTPUT_GROUP_LE                                      1024             0
 Macsec SPD                                            256             2
```

SSDPによるリソース枯渇の防止

リソースの枯渇を停止するには、最初のL3ホップとマルチキャスト状態の作成前にSSDPトラフィックを停止する必要があります。最も迅速な解決策は、このトラフィックを認識するPIMで設定されたすべてのL3インターフェイスに対して入力で適用されるIPv4アクセスコントロールリス

ト(ACL)を使用することです。「show ip mroute 239.255.255.250」コマンドを使用して確認し、各グループの「着信インターフェイス」を調べます。これは、トラフィックの送信元のL3インターフェイスを示し、複数の一意な送信元インターフェイスが存在する可能性があることに注意してください。この設定例では、SSDPがレイヤ2で動作し、L2隣接ホストがPNPサービスを検出することを許可していますが、クライアントアドバタイズメントがL3境界を越えて転送されることを防ぎ、すべてのマルチキャストルータまたはスイッチでL3マルチキャスト状態が作成されることを防ぎます。

拡張ACLを設定します。

ip access-list extended BLOCK_SSDP remark Block SSDP **deny ip any host 239.255.255.250 <-- Deny SSDP**
permit ip any any **<-- Permit any other group**

各L3インターフェイスで設定し、入力方向にACLを適用します。

Switch#**configure terminal**
Switch(config)#**interface vlan100**
Switch(config-if)#**ip access-group BLOCK_SSDP in**
Switch(config-if)#**end**